# Predicting Interest Level based on EEG Scan Data using Machine Learning Algorithms

A. Bhargava

**Abstract**—In this paper, the question, "to what extent can EEG and machine learning be used to predict the self-reported interest level of a given student" is investigated. It was found that EEG and machine learning could be used to predict the self-reported interest level of a given student to a great extent. Classifiers were able to achieve a 67% accuracy in terms of binary prediction of interest level. This was after training and validating on a set of 30 EEG scans of 10 participants who read various articles while wearing an EEG headset and then self-reported their interest level.

**Index Terms**— Brain-computer interfaces, electroencephalography, emotion recognition, machine learning

## 1 Introduction

Since the invention of Electroencephalography (EEG) technology, scientists and doctors have been attempting to use it to better understand the human brain [1]. EEG is a common, non-invasive technique for brain scanning that usually uses contact electrodes on the scalp of a subject to determine the electric potential of the scalp at that point, which reflects the neuronal activity in the underlying part of the brain [2]. However, until recently, EEG was only used for general diagnosis of brain death or seizures [3]. As computing technology progressed, scientists have begun to use machine learning algorithms to interpret more fine details from EEG scans [4]. A machine is said to be learning if it can improve its performance on a given task with more experience at said task [5]. These range from predicting memory to learning, to test performance, to name just a few [8-12]. Given these precedents, a possible next application is to determine the extent to which one can train machine learning algorithms to predict self-reported interest values in a given task based on EEG scan data.

This led to the question: To what extent can electroencephalography and machine learning be used to predict a student's self-reported interest value in a given task? The question is an important one in the realms of computer science and neuroscience because it pushes the boundaries of brain scan analysis. Without the use of relatively new machine learning algorithms, it would be practically impossible for a human to classify EEG scans in the proposed way.

In order to determine the validity of the proposed research question, it was necessary to determine whether the prediction of self-reported interest with EEG and machine learning has a basis in educational science, neuroscience, and computer science. Additionally, it needed to be established that there is, in fact, a gap in the current body of research in this area. Overall, it was found that this line of research is valid in all of these domains.

### 1.1 Background: Educational Sciences

Considering the importance of creating a classifier for brainwaves based on interest level prompts the question, is interest truly important in education? A study by Lin and Huang found that maximizing interest is of great importance for optimizing education since it has been shown that students with higher levels of interest show deeper understanding [6]. This finding was boosted by Sorić and Palekčić who found that, in the case of self-directed learning, interest is of even greater importance as a fundamental requirement for learning [7].

Given that interest has been proven to play a fundamental role in learning, it is clearly important to have a strong measurement tool for interest in order to formulate objectively beneficial educational strategies for as many students as possible, particularly for students engaging in self-learning. This study lays the groundwork for such technology that eventually may be able to be used in a classroom setting.

### 1.2 Background: Neuroscience

Comprehending the history and neuroscience behind EEG scanning was exceedingly helpful in the experimental design phase of this study, as well as in confirming that there is a gap in the current body of knowledge on the topic. This study used EEG scanning technology because fMRI[2] and other brain scanning technologies fell significantly outside of the practically non-existent research budget. According to the Encyclopaedia of Britannica, EEG scans historically have been used rarely for anything except for determining the most significant changes in neural activity [1]. For instance, the diagnosis of epileptic seizures and brain death are common applications for EEG in hospitals. Over time, EEG scan reading has been made more precise, as explained by Cohen in a more modern review of the field [2]. Cohen posits that different parts of EEG scans can correspond to general elevations in visceral

- *A. Bhargava is with Trinity College School, Port Hope, ON, Canada L1A 4K7. E-mail: abhargava@tcs.on.ca*

[2] fMRI (functional magnetic resonance imaging) is another brain scanning technique where blood flow in the brain is tracked using magnetic resonance imaging.

psychological characteristics, such as excitement or fear. For example, it has been observed that people who are agitated often have a specific oscillatory pattern in certain parts of the EEG compared to non-agitated patients.

Although it would seem as though EEG technology would eventually progress to a point at which psychological characteristics such as interest level could easily be read, findings from Burrous in his book, Standard EEG: A Research Roadmap for Neuropsychiatry, placed a limit on the level of detail that a human can detect in an EEG scan. This is because it was found that EEG can be extremely variable for healthy persons experiencing the same situations [3]. This limit was reinforced by Pernet *et al*., who stated, in their study on the efficacy of machine learning techniques for the analysis of EEG, that even signals from hospital-grade EEG are extremely noisy and generalized to broad areas of the brain when compared to other brain scanning techniques such as fMRI [4]. They came to this conclusion by performing a meta-analysis of 23 studies and meta-analyses on the use of machine learning to interpret EEG scans. Literature in the field of EEG indicated that some sort of algorithmic approach would be necessary to deduce interest from brain scan data (hence the use of machine learning).

When it came time to search for an EEG device to collect data for this study, research by Searle and Kirkup on the efficacy of various electrode types led to the decision to limit the search for EEG scanning devices to only those that used saline-soaked electrodes instead of conventional dry-contact consumer-grade electrodes in order to minimize the amount of noise in the data for this study. However, the effectiveness of consumer-grade EEG scanners was questioned heavily in the study [8].

This supposedly low quality of consumer-grade EEG scanning devices motivated the research of Maskeliunas *et al*., who decided to study which consumer-grade EEG scanning device was the best for research. They found that the Emotiv EPOC+ EEG headset was the best option for a consumer-grade EEG headset for research because of its accuracy and precision when compared to other consumer models [9]. However, all of these researchers agree that, for practically all EEGs, no matter the price bracket, there will be noise in the data. This can be caused by a great variety of factors, such as induced currents from electromagnetic waves, electric potentials from muscle contractions, or changes in electric potential caused by shifting electrodes. In any case, these apparently confounding variables, in addition to the inherent complexity of determining fine psychological characteristics based on generalized voltage potentials of neurons in various regions of the brain as measured through the skull and scalp, make it practically impossible for an unaided human to simply look at an EEG and learn which characteristics of an EEG correspond with a given independent variable (e.g. whether or not a person is interested in what they are learning), especially when the independent variable is a very specific psychological characteristic. This fact should be taken into account when examining the effectiveness of classifiers on the data.

Despite the apparent flaws with using EEG scans to predict psychological characteristics, there was still hope for predicting interest based on brain scan data. Neuroscientists and researchers are still able to interpret EEG data to predict specific psychological characteristics using computer science and mathematics to augment human processing ability. For example, Noh *et al*. utilised these machine learning algorithms to predict subsequent memory from a sample size of 18 participants with a final accuracy of 60% based on EEG scans [10]. Their method consisted of recording EEG scans of subjects attempting to memorize strings of words, letters, or numbers, and then processing the data using various machine learning algorithms. In a similar vein, a group of researchers from Boston University also attempted a study on predicting memory from EEG, this time with superior tools. They trained an algorithm that was able to accurately classify the top 5 subjects in terms of memory test performance based on EEG scan data [11]. The methodology of the two studies was extremely similar, as they recorded EEG scans of students taking memory tests, trained machine learning algorithms to correlate performance with EEG scan data and then tested it on a new set of test-takers to gauge how well their algorithm worked. Walter *et al*. from Germany took this type of research a step further and used support vector machine (SVM) learning algorithms to create an automatically optimizing arithmetic learning environment with a sample size of 13 subjects that was able to teach mathematics faster than a hard-coded alternative learning environment [12]. An SVM is a type of machine learning algorithm that is prevalent in EEG processing because of its ability to differentiate between complex data with complex decision boundaries [13]. Their approach consisted of recording EEG scans of subjects attempting to learn base-8 arithmetic, testing their learning along the way with simple tests. The researchers then trained various machine learning algorithms (including SVM's) to classify EEG scans of those who did well on tests for a particular skill and those who did not. The final learning environment used the trained algorithms to predict whether or not a student was properly absorbing the information that was being taught at a given moment. If the algorithm predicted that the student was not properly learning the content, the program would repeat the content until the algorithm predicted that the student had learned the content. Otherwise, the program would go on to the next skill.

After looking in detail at these studies, it has become apparent that the predominant algorithm was an SVM, although others were often tested simultaneously. The procedures were all generally similar, including a trial phase where subjects would complete a task with a headset on while data was being recorded. Commonly stated sources for errors include signal interference from EM waves, noise from poor electrode contact and muscle contraction, and differing mental states. These can be rectified by using an electromagnetically shielded room, ensuring proper electrode contact, and by leaving time for subjects to relax and become accustomed to wearing the EEG headset. Finally, in all studies that used dry-contact consumer-grade electrodes, it was pointed out that the sensor type was a confounding element for accuracy. This was not realistic to address in this study as a hospital or research-grade EEG far

exceeds the research budget. Overall, in terms of the neuroscience research behind the proposed study, there was firm ground and a variety of utile precedents to draw from during the design phase of this study.

### 1.3 Background: Computer Science

As stated by Michael Snyder when being interviewed about the use of artificial intelligence in medicine, "In hindsight, everything makes sense ... [T]he computers can assess even tiny differences across thousands of samples many times more accurately and rapidly than a human" [14]. Machine learning algorithms have the potential to differentiate between noise and useful signal data if properly implemented with quality data [14]. Motivated by the same reasoning that Snyder provided on the use of intelligent algorithms in medicine, there has been a large volume of recent studies that utilise machine learning and EEG to create models that predict psychological characteristics and outcomes. Any task, activity, or decision that would cause a differentiable thought process to be used appears to have a strong potential to be predicted using EEG and machine learning. Even subtle differences can be detected and utilised by machine learning algorithms to classify scans into desired categories.

An example of EEG and machine learning being used to predict psychological characteristics comes from N. H. Liu *et al*. from National Pingtung University of Science & Technology in Taiwan. They were able to create classification algorithms for determining which students were attentive and which were not based on EEG scans, which was then used to make the predictions of whether or not a student was paying attention or not based on novel EEG scans with as much as 76.82% accuracy [15].

Although there were significant limitations to the study, including poor accuracies for certain trials and noisy equipment, the experimental design was valid, and the data processing techniques were valuable. They used a Fourier transform in order to compute their data before it was fed into the learning algorithms. Fourier transforms essentially split signals into component sine and cosine wave functions, which can be extremely useful when reducing the dimensionality of time-series data. Due to the nature of the learning algorithms and techniques that were implemented in this study, dimensionality reduction is a very valuable tool.

On the other hand, P. Sajda *et al*. performed a meta-analysis of 112 studies that use machine learning algorithms and other statistical models for the classification of brain scans in various decision-making tasks. The most relevant element of the study is how the results of the analysis differed somewhat from the study by N. H. Liu *et al.* in their conclusion on the optimal way to process EEG data. Rather than using any type of complex mathematical analysis to pre-process their data, they stated that directly inputting the EEG data matrix into the chosen learning algorithm could be a valid approach [16]. However, this approach was also stated to be generally more suitable for deep learning algorithms where the algorithm conducts complex abstraction on a massive dataset at great computing cost. This would be a high-cost method of analysis as the EEG signal is sampled many times per second (commonly 128) from a minimum of 10 electrodes [2]. After only a few minutes, there would be several million values in the recorded signal. It is not effective to use this magnitude of raw data with (relatively) simple machine learning algorithms that were used in this study. For this reason, a Fourier transform was applied to the raw data in this study in order to extract more easily manipulable and usable features[3] (i.e. the frequencies that compose the EEG signal).

As described in both the neuroscience and computer science sections, there is a wealth of studies that attempt to correlate EEG data with psychological characteristics. They were unanimous about several core elements of what this type of study requires in order to increase chances of success, and the studies that were examined in the computer science section by N. H. Liu *et al.*, Maskeliunas *et al.*, and P. Sajda *et al.* had general agreement in terms of the potential errors and best practices for data collection and processing. Clearly, there was strong backing for this study.

## 3 Materials and Methods

In order to train a classification algorithm, data is required. Given that there is no pre-existing corpus of EEG data paired with interest level readings, experimental data collection was required. For this, the Emotiv EPOC+ EEG scanner was selected for this experiment. Maskeliunas *et al.* found that the Emotiv EPOC+ EEG headset was the best option for a consumer-grade EEG headset for research because of its accuracy and precision when compared to other consumer models [9].

For this experiment, a variety of subjects were needed so that the algorithm would be able to be generalized to all students (i.e. "a given student") rather than just trained for one specific person. After the institutional review board approved the experimental design, 10 subjects were used since that was within the range of the studies examined that attempt to make EEG classification algorithms. As well, using significantly more subjects would require an unrealistic amount of time to be spent on data collection. Unexpectedly, far more subjects applied to be a part of this experiment than were required, so it was possible to diversify the ages and genders of the students selected. 5 male and 5 female students were selected, with one of each in grades 9-11 and two of each in grade 12. Since this research is meant to be applicable to any "given student", it was advantageous to be able to select a reflective range of ages and genders that would be present in any high school. In the cases where there were multiple students who fit the gender and grade level requirements, the student who responded first to the invitation was selected.

A set of tasks also had to be selected to cue various interest states. A set of three readings of diverse interest levels was chosen for this task. The first reading was an excerpt from the introduction to Malcolm Gladwell's Outliers (Appendix 1.1) [17]. This was selected to generally cue people to be interested. The second reading was the introduction to the Benjamin Franklin Wikipedia page (Appendix 1.2) [18]. It was expected that this reading would have

[3] Features are processed pieces of data that are fed into algorithms (usually after dimensionality reduction)

a roughly even spread of people who were interested and people who were not. The final reading was an entire Wikipedia page on mathematics (Appendix 1.3) [19]. Given the relatively dry, academic nature of the page, and the fact that the test subjects had only been formally exposed to high school level mathematics, it was correctly assumed that this would be a generally uninteresting article. Assuming that the participants told the truth about their level of interest and that there was a relatively even and well-predicted distribution of interest levels, there ought to be a reasonable number of positive and negative cases.

The forms in appendix 2 were used to collect metadata on each participant. The order of the readings was rotated for each trial to reduce any bias in interest level resulting from the novelty of the scanner. Students were fitted with the scanner, and in accordance with the experimental design from the study by Noh *et al*., students were allowed to sit with the scanner on their heads for 2 minutes at the beginning of each trial to further reduce any bias resulting from the novelty of the scanner. They were then given the first reading for 3 minutes and 30 seconds. Once the time was up, the interest level was recorded on form C (see appendix 2.4).

Once data was collected, it was exported into .CSV format (comma separated values) for ease of use in a variety of scripting and programming languages. A Java script was coded and used to separate each trial based on timestamp markers from the CSV file (see appendix 4.1 for code). Java was used because it was the most efficient way to code a script as the experimenter had prior experience with Java coding. Following this, the Octave fast Fourier Transform (FFT) algorithm was used to extract features from the data (appendix 4.2).

The extraction of features from the raw data was necessary given the sample size of data. In machine learning, it is necessary to have fewer features[1] than training examples for the algorithm. This is because the algorithm could simply locate a single feature for each of the training examples to focus on and achieve a ~100% accuracy rating on the training data but a ~0% accuracy on new data (this is known as overfitting). A Fourier transform is an algorithm that takes in time-based data (e.g. EEG signal) and decomposes it into its component sine and cosine waves. The output of the Fourier transform is a description of the relative strength of the component sine and cosine waves that make up the original signal. This has been very useful for both machine and human analysis of machine learning algorithms in the past as it is practically impossible to simply look at EEG waveforms and tell anything meaningful about the person whether or not one is a computer or a human [2].

The Octave FFT was employed as Octave is an open source mathematical analysis language with which the researcher had previous experience. Additionally, the FFT algorithm is pre-installed with Octave. The FFT output was then averaged into 8 bands. In medical applications, the useful band frequencies range from roughly 4hz to 44hz [2], so the 8 FFT bands were split evenly from within that range. Given the differences in the relative strengths of each subject's waveforms because of varying skull/scalp/intracranial fluid conductivities, each band amplitude was measured relative to the lowest frequency band. Additionally, the log of each relative band power was taken given the extreme disparity between the lowest frequency power and the rest. This resulted in a reasonably well-scaled set of features.

Features scaling is important in machine learning because of the method by which algorithms optimize their parameters to best predict the desired output (in this case interest level) based on the input. To understand this, consider the example of an algorithm meant to predict housing prices based on square footage. This algorithm has only one parameter, which is the number by which it multiplies the inputted square footage to predict the price of the house. The algorithm can be mathematically encoded as follows:

$$P_m(x) = mx$$

*Where P(x) is the predicted price of the house based on square footage, m is the multiplier parameter, and x is the input (square footage).*

In order to learn from training data, the algorithm's 'goal' is to find the best value for m for accurately predicting the price of a house.

Consider also the function that calculates a number for how good the algorithm, $P_m(x)$, is at predicting housing prices:

$$C(P_m) = \sum_{i=0}^{n} P_m(X_i) - Y_i$$

*Where X and Y are matrices for which the housing price for the house with square footage Xi is equal to Yi, and Σi is summing from all values i = 0 to i = n where n is the length of X and Y.*

Essentially, if $P_m$ is able to predict the corresponding $Y_i$ to $X_i$ very well (i.e. the difference between $P_m(X_i)$ and $Y_i$ is very small), $C(P_m)$ will be very small, indicating that the algorithm $P_m$ is very effective.

When plotted, the graph of $C(P_m)$ vs. m should look like the letter 'U'. There is some optimal value of m where Pm predicts housing prices well. As m deviates from that value, $C(P_m)$ rises, leading to the 'U' shape of the function. The training objective is therefore to find the value of m at the bottom of that 'U' shaped function.

To do this, the algorithm will employ an optimization technique called gradient descent. Essentially, the algorithm will start at a random point on the curve and 'look around' to see where the natural slope of the curve leads. It then takes a 'step' in that direction and 'looks around' again to see where the natural slope of the curve leads. The exact mechanism by which the algorithm does this involves rather complex multivariable calculus, so it will not be described herein. After repeating these steps enough

[1] Features are processed pieces of data that are fed into algorithms (usually after dimensionality reduction).

times, the algorithm will arrive at a point where there is a local optimum, which can be thought of as the bottom of the U.

When there is only one parameter (and therefore one non-C($P_m$) dimension), the scale of the data that is being trained on does not matter. However, things become more complicated as the number of dimensions increases. For instance, if one wanted to have another input, for example a house's proximity to the nearest body of water, the cost function gains another dimension and also begins to resemble a bowl. There is still an optimal point, but there are two dimensions to contend with. To employ gradient descent on this new function, the algorithm alternates between taking steps in one dimension and then the other until it reaches a local optimum.

An issue arises when the scale of one dimension is radically different than the other. Instead of looking like an even bowl, the C(Pm) function will begin to look more like a canoe. The algorithm is now likely to overshoot or undershoot in one of the dimensions as it will be taking (roughly) equal steps in each dimension, but in one of the dimensions the function is much more or less steep. The further basis of this problem is rooted in the exact mathematical method by which gradient descent is conducted.

This problem is compounded as more and more dimensions are added. In the cases of this study, there are eight total dimensions (one for each frequency band from the EEG). Therefore, some sort of scaling must occur to ensure that a local optimum is found well. For this reason, the data was processed in the way described above [17].

These band values were then concatenated into a single spreadsheet and the interest level for each set of bands was added as another column. The data was uploaded to the Azure Machine Learning platform, and a variety of machine learning algorithms were applied, using the band power columns as features and the binary interest level column as labels.

The Microsoft Azure Machine Learning library was used in this study for final data processing. Although the hard-coding of the machine learning algorithms in a language like Octave was an option, a machine learning library was deemed preferable because of the fact that the implementation of the algorithms is optimized by qualified people, meaning that no further benefit would be obtained via the hand-coding of the algorithms. Azure was chosen because it is a free platform with a wide variety of algorithms ready to be used. As well, the graphical user interface makes it extremely easy to use. This being said, the same results should be found using any other library or resource for the same machine learning algorithms with the same hyperparameters.

A variety of machine learning algorithms were tried and tested for the data, including an SVM, an artificial neural network, a random forest algorithm, and a logistic regression classifier. These were chosen because of their prevalence in the surrounding literature pertaining to the classification of EEG data, and because they are commonly used for data of this type (i.e. training examples with less than 102 features and one binary label). The complex mechanisms by which these algorithms function will not be described herein, although further reading can be done in Stanford Professor and machine learning pioneer Andrew Ng's open source course [5].

Usually, the subset of the data upon which a learning algorithm is trained (the training set) is comprised of 70% of the original data and the set of data upon which the accuracy of the classifier is tested (the independent holdout set) is the remaining 30%. This is done so that it can be tested (using the independent holdout set) whether or not the algorithm is truly learning about the data instead of just "memorizing" the correct output for each input that it was trained on (known as overfitting) [5]. However, this approach of simply partitioning the data into a training set and an independent holdout set can be problematic, especially with smaller datasets, since the algorithm is essentially 'wasting' 30% of the data during the training phase and 70% during the validation phase. A technique called K-fold cross-validation addresses this problem by splitting the data into K-number folds (in this case 10) and training the algorithm on K-1 of the folds and testing it on the remaining fold. It does this K times, with a different training and testing subset each time. In the end, the accuracy of the model is determined based on the average accuracy on each fold. By this method, the model trained and validated on all of the data without training and validating on the same data at the same time, without 'wasting' any of the data.

## 4 Results

The accuracy of each classification algorithm is shown in table 1. As stated in the Materials and Methods section, the algorithms were trained on a subset of the collected data consisting of 22 brain scans, with half interested and half uninterested scans used. The highest accuracy was achieved using an SVM algorithm, for which a final classification accuracy of 70% was obtained when 10-fold cross-validation was implemented. The lambda value for the SVM was set to 0.001, and the rest of the hyperparameters can be seen in table 2. Given the complex method by which machine learning algorithms such as SVM's optimize their parameters and classify inputs, the meaning of each hyperparameter will not be described herein. The hyperparameters for the rest of the algorithms can be seen in appendix 3.

| Algorithm | Classification Accuracy |
|---|---|
| SVM (Support Vector Machine) | 70% |
| Logistic Regression | 67% |
| Artificial Neural Network | 61% |
| Decision Forest | 60% |

*Table 1: The accuracy of each algorithm in 10-fold cross-validation.*

**Support Vector Machine Classifier**

**Settings**

| Setting | Value |
|---|---|
| Lambda | 0.001 |
| Num Iterations | 1 |
| Normalize Features | True |
| Perform Projection | False |
| Allow Unknown Levels | True |
| Random Number Seed | |

*Table 2: SVM hyperparameters.*

## 5 Discussion

The results indicate that the interest of a given student can be predicted to a great extent using EEG and machine learning. An accuracy of 66% is high relative to other papers that attempt to create classifiers for EEG data and other similar psychological characteristics [10]-[12] and is especially notable considering the limitations in this study. This high accuracy reinforces the findings of Noh *et al.*, Matzen *et al.*, and Walter *et al.* that attempted to classify EEG scan data based on psychological characteristics.

### 5.1 Importance

This was an important investigation for a variety of reasons. The prediction of psychological characteristics based on brain scans has a tremendous number of applications. In this case, the most immediate application is in education. At present, a teacher relies on visual cues and student performance to gauge interest level, and it is difficult for any person to determine this characteristic reliably and accurately [6]. The use of technology to make this more precise and accurate will give teachers, both human and artificial, yet another tool to use when optimizing an educational strategy for a student. Given that interest has been proven to be an important psychological characteristic for education [7], this research is important as it adds to teachers' toolbox for educating the next generation.

### 5.2 Limitations

In this study, there are plenty of limitations to consider. The two main limitations were the noise in the data and the volume of data.

As stated in the materials and methods section, the headset used is a consumer-grade headset. Its use of saline-saturated felt pads instead of the conductive gels used in hospital-grade EEG's adds noise as the pads could easily be dislodged and the conductivity was not as reliable. Additionally, EEG in general has been shown to be a noisy brain scanning technique when compared to techniques such as fMRI [1]. This noise inhibits attempts to classify the data as the core data is hidden behind a veil of noise.

The volume of data is also a limitation. Although it is common practice to use less than 30 test subjects in studies attempting to correlate EEG data with psychological characteristics, the use of just 10 participants was a hindrance when it came time to process the data. Each participant read 3 articles, so there were 3 data points per participant. However, some needed to be removed to balance the dataset so that there were half interested reading and half uninterested readings, resulting in only 22 available data points to both train and validate from. Given that machine learning relies on a large volume of data to identify patterns and trends in a given domain, it is a significant problem to have so little data. The lack of training data casts the findings into some doubt as it could be simply random chance that the classifiers worked the way that they did. In any case, it would certainly be beneficial to have more data.

### 5.3 Next Steps

There are a wide variety of paths that future research could take in this field. For example, this study attempted to create a generalized classification algorithm for any person's brainwaves. However, given the natural variation in human brainwaves described by J. Cohen in his book, Electroencephalography, it may be more useful and feasible to create personalized classifiers for each subject [2]. This would require significantly more data per subject, but it would reflect more accurately how the algorithms would work if they were to be implemented in a practical educational context, as discussed in the Importance section. The algorithms that predict the interest of each student could continue to train on the new brain data that they obtain from each student for maximum predictive power. This individualized continual-learning model would help to elucidate the true extent to which a student's interest level can be predicted by EEG and machine learning algorithms.

Research should also be expanded by investigating the prediction of other useful characteristics based on EEG. For example, one could use these same techniques to investigate if there is a neurological difference between two teaching methods, or if there is a neurological difference between those with different 'learning styles'. This could even be used to investigate the neurological roots of a person's response to authority. There is a practically limitless number of possible paths of investigation with this technology and technique.

Additionally, further research and optimization should be pursued in the realm of feature extraction from EEG. Implementing a Fourier transform to extract features does not take into account shifts in waves and frequencies. Given the wealth of time-frequency analysis[4] techniques that have been invented, it would be extremely useful to

[4] Techniques that study a time-bound data in terms of time and frequency at the same time

learn more about which one(s) work the best for EEG classification in various contexts.

Finally, the research could be validated by repeating this same experiment except with superior EEG equipment, such as with a hospital-grade EEG machine, and with a larger sample size.

## 7 Conclusion

In summary, it was found that it is, in fact, possible to create and train a machine learning classifier to predict a student's interest level with significantly greater than random accuracy. Though there are plenty of improvements to be made to the experiment and plenty of future experimental paths, the results of this experiment are promising.

## Acknowledgment

The author wishes to thank Trinity College School. This work was supported in part by a grant from Trinity College School (used for the purchase of accompanying EEG software).

## References

[1] S. Singh and K. Rogers, "Electroencephalography," in Encyclopædia Britannica: Physiology, The Editors of Encyclopædia Britannica: Encyclopædia Britannica, Inc, 2017 [Online]. Available: Britannica.com

[2] J. Cohen, "Electroencephalography," AccessScience, 2014. [Online]. Available: www.accessscience.com. [Accessed November 6, 2017].

[3] N. Burrous, Standard EEG: A Research Roadmap for Neuropsychiatry. New York, NY: Springer, 2013, pp. 15-24.

[4] C. R. Pernet *et al.*, "Single-trial analyses: Why bother?," Frontiers in Psychology, vol. 2, Article 322, November 2011.

[5] A. Ng, Class Lecture, Topic: "Introduction to Machine Learning" Machine Learning, Coursera, 2017.

[6] S. H. Lin and Y. C. Huang, "Examining charisma in relation to students' interest in learning," Active Learning in Higher Education, vol. 17(2) 139-151, 2016. [Online]. Available: IEEE Xplore, http://www.ieee.org. [Accessed Sept. 10, 2010].

[7] I. Sorić and M. Palekčić. "The role of students' interests in self-regulated learning: The relationship between students' interests, learning strategies and causal attributions," European Journal of Psychology of Education, vol. 24, no. 4, pp. 545-565, 2009.

[8] A. Searle and L. Kirkup, "A direct comparison of wet, dry and insulating bioelectric recording electrodes," Psychological Measurement, vol. 23, no. 2, 1999. [Online]. Available: IOP Science, http://iopscience.iop.org. [Accessed Nov. 10, 2017].

[9] R. Maskeliunas *et al*. "Consumer-grade EEG devices: are they usable for control tasks?," PeerJ, vol. 4, March, 2016. doi: 10.7717/peerj.1746.

[10] E. Noh *et al.*, "Using Single-trial EEG to Predict and Analyze Subsequent Memory," Neuroimage, vol. 84, no. 1, 2014. [Online]. Available: NCBI, https://www.ncbi.nlm.nih.gov. [Accessed Nov. 10, 2017].

[11] L. Matzen *et al.*, "Monitoring Brain Activity During Studying to Predict Test Performance," ScienceDaily, 2012. [Online]. Available: https://www.sciencedaily.com. [Accessed: Nov. 10, 2017].

[12] C. Walter *et al.*, "Online EEG-Based Workload Adaptation of an Arithmetic Learning Environment," Frontiers in Human Medical Sciences, vol. 11, Article 286, May 2017.

[13] A. Ng, Class Lecture, Topic: "SVM's" Machine Learning, Coursera, 2017.

[14] K. Conger, "Computers trounce pathologists in predicting lung cancer type, severity," Stanford Medicine News Center, para. 11, August 16, 2016. [Online]. Available: https://med.stanford.edu. [Accessed Nov. 10, 2017].

[15] N. H. Liu *et al.*, "Recognizing the Degree of Human Attention Using EEG Signals from Mobile Sensors," Sensors, vol. 13, no. 8, p. 10273-10286, 2013. [Online]. Available: MDPI, http://www.mdpi.com. [Accessed: Nov 10, 2017].

[16] P. Sajda *et al.*, "Single-Trial Analysis of Neuroimaging Data: Inferring Neural Networks Underlying Perceptual Decision-Making in the Human Brain," IEEE Rev Biomed Eng., 2009, no. 2, p. 97–109.

[17] M. Gladwell, *Outliers: The Story of Success*. New York, NY: Little, Brown and Company, 2008.

[18] "Benjamin Franklin," in Wikipedia. Wikimedia Foundation, [online document], 2018. Available: Wikipedia, https://www.wikipedia.org [Accessed: January, 2018].

[19] "Mathematics," in Wikipedia. Wikimedia Foundation, [online document], 2018. Available: Wikipedia, https://www.wikipedia.org [Accessed: January, 2018].

[20] A. Ng, Class Lecture, Topic: "Solving Overfitting in Cost Function" Machine Learning, Coursera, 2017.

**A. Bhargava** has taken no academic degrees thus far, although he is set to receive a Canadian High School Diploma from Trinity College School in June of 2018. He has been employed at Maine Doctor's Office, IMCare, and Fluent.AI, in addition to his own freelance development business. He has yet to be associated with any journals or conferences. Bhargava has yet to achieve major professional or academic honours. His research interests are broad and somewhat undefined, but generally include computer scientific, medical, physics, biological, and engineering-based research.

## Appendix 1: Readings

### 1.1 Outliers Excerpt (Malcolm Gladwell) [17]

**Outliers: The Story of Success - Malcolm Gladwell**

For almost a generation, psychologists around the world have been engaged in a spirited debate over a question that most of us would consider to have been settled years ago. The question is this: is there such a thing as innate talent? The obvious answer is yes. Not every hockey player born in January ends up playing at the professional level. Only some do--the innately talented ones. Achievement is talent plus preparation. The problem with this view is that the closer psychologists look at the careers of the gifted, the smaller the role innate talent seems to play and the bigger the role preparation seems to play.

Exhibit A in the talent argument is a study done in the early 1990s by the psychologist K. Anders Ericsson and two colleagues at Berlin's elite Academy of Music. With the help of the Academy's professors, they divided the school's violinists into three groups. In the first group were the stars, the students with the potential to become world-class soloists. In the second were those judged to be merely "good." In the third were students who were unlikely to ever play professionally and who intended to be music teachers in the public school system. All of the violinists were then asked the same question: over the course of your entire career, ever since you first picked up the violin, how many hours have you practiced?

Everyone from all three groups started playing at roughly the same age, around five years old. In those first few years, everyone practiced roughly the same amount, about two or three hours a week. But when the students were around the age of eight, real differences started to emerge. The students who would end up the best in their class began to practice more than everyone else: six hours a week by age nine, eight hours a week by age twelve, sixteen hours a week by age fourteen, and up and up, until by the age of twenty they were practicing--that is, purposefully and single-mindedly playing their instruments with the intent to get better--well over thirty hours a week. In fact, by the age of twenty, the elite performers had each totaled ten thousand hours of practice. By contrast, the merely good students had totaled eight thousand hours, and the future music teachers had totaled just over four thousand hours.

Ericsson and his colleagues then compared amateur pianists with professional pianists. The same pattern emerged. The amateurs never practiced more than about three hours a week over the course of their childhood, and by the age of twenty they had totaled two thousand hours of practice. The professionals, on the other hand, steadily increased their practice time every year, until by the age of twenty they, like the violinists, had reached ten thousand hours.

The striking thing about Ericsson's study is that he and his colleagues couldn't find any "naturals," musicians who floated effortlessly to the top while practicing a fraction of the time their peers did. Nor could they find any "grinds," people who worked harder than everyone else, yet just didn't have what it takes to break the top ranks. Their research suggests that once a musician has enough ability to get into a top music school, the thing that distinguishes one performer from another is how hard he or she works. That's it. And what's more, the people at the very top don't work just harder or even much harder than everyone else. They work much, much harder.

The idea that excellence at performing a complex task requires a critical minimum level of practice surfaces again and again in studies of expertise. In fact, researchers have settled on what they believe is the magic number for true expertise: ten thousand hours.

"The emerging picture from such studies is that ten thousand hours of practice is required to achieve the level of mastery associated with being a world-class expert--in anything," writes the neurologist Daniel Levitin. "In study after study, of composers, basketball players, fiction writers, ice skaters, concert pianists, chess players, master criminals, and what have you, this number comes up again and again. Of course, this doesn't address why some people get more out of their practice sessions than others do. But no one has yet found a case in which true world- class expertise was accomplished in less time. It seems that it takes the brain this long to assimilate all that it needs to know to achieve true mastery."

This is true even of people we think of as prodigies. Mozart, for example, famously started writing music at six. But, writes the psychologist Michael Howe in his book Genius Explained, by the standards of mature composers, Mozart's early works are not outstanding. The earliest pieces were all probably written down by his father, and perhaps improved in the process. Many of Wolfgang's childhood compositions, such as the first seven of his concertos for piano and orchestra, are largely arrangements of works by other composers. Of those concertos that only contain music original to Mozart, the earliest that is now regarded as a masterwork (No. 9, K. 271) was not com posed until he was twenty-one: by that time Mozart had already been composing concertos for ten years.

The music critic Harold Schonberg goes further: Mozart, he argues, actually "developed late," since he didn't produce his greatest work until he had been composing for more than twenty years.

To become a chess grandmaster also seems to take about ten years. (Only the legendary Bobby Fischer got to that elite level in less than that amount of time: it took him nine years.) And what's ten years? Well, it's roughly how long it takes to put in ten thousand hours of hard practice. Ten thousand hours is the magic number of greatness.

Here is the explanation for what was so puzzling about the rosters of the Czech and Canadian national sports teams. There was practically no one on those teams born after September 1, which doesn't seem to make any sense. You'd think that there should be a fair number of Czech hockey or soccer prodigies born late in the year who are so talented that they eventually make their way into the top tier as young adults, despite their birth dates.

But to Ericsson and those who argue against the primacy of talent, that isn't surprising at all. That late-born prodigy doesn't get chosen for the all-star team as an eight-

year-old because he's too small. So he doesn't get the extra practice. And without that extra practice, he has no chance at hitting ten thousand hours by the time the professional hockey teams start looking for players. And without ten thousand hours under his belt, there is no way he can ever master the skills necessary to play at the top level. Even Mozart--the greatest musical prodigy of all time--couldn't hit his stride until he had his ten thousand hours in. Practice isn't the thing you do once you're good. It's the thing you do that makes you good.

The other interesting thing about that ten thousand hours, of course, is that ten thousand hours is an enormous amount of time. It's all but impossible to reach that number all by yourself by the time you're a young adult. You have to have parents who encourage and support you. You can't be poor, because if you have to hold down a part-time job on the side to help make ends meet, there won't be time left in the day to practice enough. In fact, most people can reach that number only if they get into some kind of special program--like a hockey all-star squad--or if they get some kind of extraordinary opportunity that gives them a chance to put in those hours.

## 1.2 Wikipedia Mathematics Introduction [19]

**Mathematics - Wikipedia, 2018**

Mathematics (from Greek μάθημα máthēma, "knowledge, study, learning") is the study of topics such as quantity (numbers),[1] structure,[2] space,[1] and change.[3][4][5] There are many views among mathematicians and philosophers as to the exact scope and definition of mathematics.[6][7]

Mathematicians seek out patterns[8][9] and use them to formulate new conjectures. Mathematicians resolve the truth or falsity of conjectures by mathematical proof. When mathematical structures are good models of real phenomena, then mathematical reasoning can provide insight or predictions about nature. Through the use of abstraction and logic, mathematics developed from counting, calculation, measurement, and the systematic study of the shapes and motions of physical objects. Practical mathematics has been a human activity from as far back as written records exist. The research required to solve mathematical problems can take years or even centuries of sustained inquiry.

Rigorous arguments first appeared in Greek mathematics, most notably in Euclid's Elements. Since the pioneering work of Giuseppe Peano (1858–1932), David Hilbert (1862–1943), and others on axiomatic systems in the late 19th century, it has become customary to view mathematical research as establishing truth by rigorous deduction from appropriately chosen axioms and definitions. Mathematics developed at a relatively slow pace until the Renaissance, when mathematical innovations interacting with new scientific discoveries led to a rapid increase in the rate of mathematical discovery that has continued to the present day.[10]

Galileo Galilei (1564–1642) said, "The universe cannot be read until we have learned the language and become familiar with the characters in which it is written. It is written in mathematical language, and the letters are triangles, circles and other geometrical figures, without which means it is humanly impossible to comprehend a single word. Without these, one is wandering about in a dark labyrinth."[11] Carl Friedrich Gauss (1777–1855) referred to mathematics as "the Queen of the Sciences".[12] Benjamin Peirce (1809–1880) called mathematics "the science that draws necessary conclusions".[13] David Hilbert said of mathematics: "We are not speaking here of arbitrariness in any sense. Mathematics is not like a game whose tasks are determined by arbitrarily stipulated rules. Rather, it is a conceptual system possessing internal necessity that can only be so and by no means otherwise."[14] Albert Einstein (1879–1955) stated that "as far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality."[15]

Mathematics is essential in many fields, including natural science, engineering, medicine, finance and the social sciences. Applied mathematics has led to entirely new mathematical disciplines, such as statistics and game theory. Mathematicians also engage in pure mathematics, or mathematics for its own sake, without having any application in mind. There is no clear line separating pure and applied mathematics, and practical applications for what began as pure mathematics are often discovered.[16]

**History**

The history of mathematics can be seen as an ever-increasing series of abstractions. The first abstraction, which is shared by many animals,[17] was probably that of numbers: the realization that a collection of two apples and a collection of two oranges (for example) have something in common, namely quantity of their members.

Greek mathematician Pythagoras (c. 570 BC – c. 495 BC), commonly credited with discovering the Pythagorean theorem

Greek mathematician Pythagoras (c. 570 BC – c. 495 BC), commonly credited with discovering the Pythagorean theorem

Mayan numerals

Mayan numerals

As evidenced by tallies found on bone, in addition to recognizing how to count physical objects, prehistoric peoples may have also recognized how to count abstract quantities, like time – days, seasons, years.[18]

Evidence for more complex mathematics does not appear until around 3000 BC, when the Babylonians and Egyptians began using arithmetic, algebra and geometry for taxation and other financial calculations, for building and construction, and for astronomy.[19] The earliest uses of mathematics were in trading, land measurement, painting and weaving patterns and the recording of time.

In Babylonian mathematics, elementary arithmetic (addition, subtraction, multiplication and division) first appears in the archaeological record. Numeracy pre-dated writing and numeral systems have been many and diverse, with the first known written numerals created by Egyptians in Middle Kingdom texts such as the Rhind Mathematical Papyrus.[citation needed]

Between 600 and 300 BC the Ancient Greeks began a systematic study of mathematics in its own right with Greek mathematics.[20]

Persian mathematician Al-Khwarizmi (c. 780 – c. 850), the inventor of algebra.
Persian mathematician Al-Khwarizmi (c. 780 – c. 850), the inventor of algebra.
During the Golden Age of Islam, especially during the 9th and 10th centuries, mathematics saw many important innovations building on Greek mathematics: most of them include the contributions from Persian mathematicians such as Al-Khwarismi, Omar Khayyam and Sharaf al-Dīn al-Ṭūsī.

Mathematics has since been greatly extended, and there has been a fruitful interaction between mathematics and science, to the benefit of both. Mathematical discoveries continue to be made today. According to Mikhail B. Sevryuk, in the January 2006 issue of the Bulletin of the American Mathematical Society, "The number of papers and books included in the Mathematical Reviews database since 1940 (the first year of operation of MR) is now more than 1.9 million, and more than 75 thousand items are added to the database each year. The overwhelming majority of works in this ocean contain new mathematical theorems and their proofs."[21]

**Etymology**
The word mathematics comes from Ancient Greek μάθημα (máthēma), meaning "that which is learnt",[22] "what one gets to know", hence also "study" and "science", and in modern Greek just "lesson". The word máthēma is derived from μανθάνω (manthano), while the modern Greek equivalent is μαθαίνω (mathaino), both of which mean "to learn". In Greece, the word for "mathematics" came to have the narrower and more technical meaning "mathematical study" even in Classical times.[23] Its adjective is μαθηματικός (mathēmatikós), meaning "related to learning" or "studious", which likewise further came to mean "mathematical". In particular, μαθηματικὴ τέχνη (mathēmatikḗ tékhnē), Latin: ars mathematica, meant "the mathematical art".

Similarly, one of the two main schools of thought in Pythagoreanism was known as the mathēmatikoi (μαθηματικοί)—which at the time meant "teachers" rather than "mathematicians" in the modern sense.

In Latin, and in English until around 1700, the term mathematics more commonly meant "astrology" (or sometimes "astronomy") rather than "mathematics"; the meaning gradually changed to its present one from about 1500 to 1800. This has resulted in several mistranslations: a particularly notorious one is Saint Augustine's warning that Christians should beware of mathematici, addressing astrologers by this notion, which is sometimes misinterpreted as a condemnation of mathematicians.[24]

The apparent plural form in English, like the French plural form les mathématiques (and the less commonly used singular derivative la mathématique), goes back to the Latin neuter plural mathematica (Cicero), based on the Greek plural τα μαθηματικά (ta mathēmatiká), used by Aristotle (384–322 BC), and meaning roughly "all things mathematical"; although it is plausible that English borrowed only the adjective mathematic(al) and formed the noun mathematics anew, after the pattern of physics and metaphysics, which were inherited from Greek.[25] In English, the noun mathematics takes singular verb forms. It is often shortened to maths or, in English-speaking North America, math.[26]

## 1.3 Wikipedia Benjamin Franklin Introduction [18]

**Benjamin Franklin's Beginnings - Wikipedia, 2018**

**Philadelphia**
At age 17, Benjamin Franklin ran away to Philadelphia, Pennsylvania, seeking a new start in a new city. When he first arrived, he worked in several printer shops around town, but he was not satisfied by the immediate prospects. After a few months, while working in a printing house, Franklin was convinced by Pennsylvania Governor Sir William Keith to go to London, ostensibly to acquire the equipment necessary for establishing another newspaper in Philadelphia. Finding Keith's promises of backing a newspaper empty, Franklin worked as a typesetter in a printer's shop in what is now the Church of St Bartholomew-the-Great in the Smithfield area of London. Following this, he returned to Philadelphia in 1726 with the help of Thomas Denham, a merchant who employed Franklin as clerk, shopkeeper, and bookkeeper in his business.[14]

**Junto and Library**
In 1727, Benjamin Franklin, then 21, created the Junto, a group of "like minded aspiring artisans and tradesmen who hoped to improve themselves while they improved their community." The Junto was a discussion group for issues of the day; it subsequently gave rise to many organizations in Philadelphia.[15] The Junto was modeled after English coffeehouses that Franklin knew well, and which had become the center of the spread of Enlightenment ideas in Britain.[16][17]

Reading was a great pastime of the Junto, but books were rare and expensive. The members created a library initially assembled from their own books after Franklin wrote:

A proposition was made by me that since our books were often referr'd to in our disquisitions upon the inquiries, it might be convenient for us to have them altogether where we met, that upon occasion they might be consulted; and by thus clubbing our books to a common library, we should, while we lik'd to keep them together, have each of us the advantage of using the books of all the other members, which would be nearly as beneficial as if each owned the whole.[18]

This did not suffice, however. Franklin conceived the idea of a subscription library, which would pool the funds of the members to buy books for all to read. This was the birth of the Library Company of Philadelphia: its charter was composed by Franklin in 1731. In 1732, Franklin hired the first American librarian, Louis Timothee. The Library Company is now a great scholarly and research library.[19]

**Newspaperman**
Benjamin Franklin (center) at work on a printing press. Reproduction of a Charles Mills painting by the Detroit Publishing Company.
Benjamin Franklin (center) at work on a printing press. Reproduction of a Charles Mills painting by the Detroit Publishing Company.
Upon Denham's death, Franklin returned to his former trade. In 1728, Franklin had set up a printing house in partnership with Hugh Meredith; the following year he became the publisher of a newspaper called The Pennsylvania Gazette. The Gazette gave Franklin a forum for agitation about a variety of local reforms and initiatives through printed essays and observations. Over time, his commentary, and his adroit cultivation of a positive image as an industrious and intellectual young man, earned him a great deal of social respect. But even after Franklin had achieved fame as a scientist and statesman, he habitually signed his letters with the unpretentious 'B. Franklin, Printer.'[14]

In 1732, Ben Franklin published the first German-language newspaper in America – Die Philadelphische Zeitung – although it failed after only one year, because four other newly founded German papers quickly dominated the newspaper market.[20] Franklin printed Moravian religious books in German. Franklin often visited Bethlehem, Pennsylvania staying at the Moravian Sun Inn.[21] In a 1751 pamphlet on demographic growth and its implications for the colonies, he called the Pennsylvania Germans "Palatine Boors" who could never acquire the "Complexion" of the English settlers and to "Blacks and Tawneys" as weakening the social structure of the colonies. Although Franklin apparently reconsidered shortly thereafter, and the phrases were omitted from all later printings of the pamphlet, his views may have played a role in his political defeat in 1764.[22]

Franklin saw the printing press as a device to instruct colonial Americans in moral virtue. In Benjamin Franklin's Journalism, Ralph Frasca argues he saw this as a service to God, because he understood moral virtue in terms of actions, thus, doing good provides a service to God. Despite his own moral lapses, Franklin saw himself as uniquely qualified to instruct Americans in morality. He tried to influence American moral life through construction of a printing network based on a chain of partnerships from the Carolinas to New England. Franklin thereby invented the first newspaper chain. It was more than a business venture, for like many publishers since, he believed that the press had a public-service duty.[23]

When Franklin established himself in Philadelphia, shortly before 1730, the town boasted two "wretched little" news sheets, Andrew Bradford's The American Weekly Mercury, and Samuel Keimer's Universal Instructor in all Arts and Sciences, and Pennsylvania Gazette. This instruction in all arts and sciences consisted of weekly extracts from Chambers's Universal Dictionary. Franklin quickly did away with all this when he took over the Instructor and made it The Pennsylvania Gazette. The Gazette soon became Franklin's characteristic organ, which he freely used for satire, for the play of his wit, even for sheer excess of mischief or of fun. From the first, he had a way of adapting his models to his own uses. The series of essays called "The Busy-Body", which he wrote for Bradford's American Mercury in 1729, followed the general Addisonian form, already modified to suit homelier conditions. The thrifty Patience, in her busy little shop, complaining of the useless visitors who waste her valuable time, is related to the ladies who address Mr. Spectator. The Busy-Body himself is a true Censor Morum, as Isaac Bickerstaff had been in the Tatler. And a number of the fictitious characters, Ridentius, Eugenius, Cato, and Cretico, represent traditional 18th-century classicism. Even this Franklin could use for contemporary satire, since Cretico, the "sowre Philosopher", is evidently a portrait of Franklin's rival, Samuel Keimer.[citation needed]

As time went on, Franklin depended less on his literary conventions, and more on his own native humor. In this there is a new spirit—not suggested to him by the fine breeding of Addison, or the bitter irony of Swift, or the stinging completeness of Pope. The brilliant little pieces Franklin wrote for his Pennsylvania Gazette have an imperishable place in American literature.[citation needed]

The Pennsylvania Gazette, like most other newspapers of the period, was often poorly printed. Franklin was busy with a hundred matters outside of his printing office, and never seriously attempted to raise the mechanical standards of his trade. Nor did he ever properly edit or collate the chance medley of stale items that passed for news in the Gazette. His influence on the practical side of journalism was minimal.[citation needed] On the other hand, his advertisements of books show his very great interest in popularizing secular literature. Undoubtedly his paper contributed to the broader culture that distinguished Pennsylvania from her neighbors before the Revolution. Like many publishers, Franklin built up a book shop in his printing office; he took the opportunity to read new books before selling them.[citation needed]

Franklin had mixed success in his plan to establish an inter-colonial network of newspapers that would produce a profit for him and disseminate virtue.[24] He began in Charleston, South Carolina, in 1731. After the second editor died, his widow Elizabeth Timothy took over and made it a success, 1738–46. She was one of the colonial era's first woman printers.[25] For three decades Franklin maintained a close business relationship with her and her son Peter who took over in 1746.[26] The Gazette had a policy of impartiality in political debates, while creating the opportunity for public debate, which encouraged others to challenge authority. Editor Peter Timothy avoided blandness and crude bias, and after 1765 increasingly took a patriotic stand in the growing crisis with Great Britain.[27] However, Franklin's Connecticut Gazette (1755–68) proved unsuccessful.[28]

**Freemason**
In 1731, Franklin was initiated into the local Masonic lodge. He became Grand Master in 1734, indicating his rapid rise to prominence in Pennsylvania.[29][30] That same year, he edited and published the first Masonic book in the Americas, a reprint of James Anderson's Constitutions of the Free-Masons. Franklin remained a Freemason for the rest of his life.[31][32]

**Common-law marriage to Deborah Read**

At age 17 in 1723, Franklin proposed to 15-year-old Deborah Read while a boarder in the Read home. At that time, Read's mother was wary of allowing her young daughter to marry Franklin, who was on his way to London at Governor Sir William Keith's request, and also because of his financial instability. Her own husband had recently died, and she declined Franklin's request to marry her daughter.[14]

While Franklin was in London, his trip was extended, and there were problems with Sir William's promises of support. Perhaps because of the circumstances of this delay, Deborah married a man named John Rodgers. This proved to be a regrettable decision. Rodgers shortly avoided his debts and prosecution by fleeing to Barbados with her dowry, leaving her behind. Rodgers's fate was unknown, and because of bigamy laws, Deborah was not free to remarry.

Franklin established a common-law marriage with Deborah Read on September 1, 1730. They took in Franklin's recently acknowledged young illegitimate son William and raised him in their household. They had two children together. Their son, Francis Folger Franklin, was born in October 1732 and died of smallpox in 1736. Their daughter, Sarah "Sally" Franklin, was born in 1743 and grew up to marry Richard Bache, have seven children, and look after her father in his old age.

Deborah's fear of the sea meant that she never accompanied Franklin on any of his extended trips to Europe, and another possible reason why they spent so much time apart is that he may have blamed her for preventing their son Francis from being vaccinated against the disease that subsequently killed him.[33] Deborah wrote to him in November 1769 saying she was ill due to "dissatisfied distress" from his prolonged absence, but he did not return until his business was done.[34] Deborah Read Franklin died of a stroke in 1774, while Franklin was on an extended mission to England; he returned in 1775.

## APPENDIX 2: EXPERIMENT FORMS

### 2.1 Consent Form

Consent Form - Subject #__
Aman Bhargava's AP Capstone Research Project: To what extent can self-reported interest be predicted using EEG and machine learning?

---

I hereby consent to having my EEG scan recorded and used anonymously for the duration of this study that is to be conducted by Aman Bhargava for his AP Capstone research project. I also consent to having my answers to the debriefing questions used anonymously in aforementioned study, and I understand the following negligible risks associated with the use of a consumer-grade EEG headset, namely the risk of static electric shock (comparable to touching any USB-connected device) and any irritation from felt pads hydrated with salt water.

Print Name: ____________________________________
Date: ____________________________________
Signature: ____________________________________

### 2.2 Subject Background Information

Form A: Background Information - Subject #__
Aman Bhargava's AP Capstone Research Project: To what extent can self-reported interest be predicted using EEG and machine learning?

1. Age: __
2. Grade Level: __
3. What is your gender? *Male Female*
4. On a scale of 1-5, how easily distracted are you? *Not easily* 1 2 3 4 5 *Very easily distracted*
5. Do you have a medical history of seizures? *Yes No*
6. Have you been diagnosed with any learning disabilities? *Yes No*
7. Did you use conditioner when you last bathed? *Yes No*
8. Are you currently on any medication? *Yes No*
9. On a scale of 1-5, how hungry are you feeling? *Not at all* 1 2 3 4 5 *Very hungry*
10. Did you consume any caffeine during the last 5 hours? *Yes No*

[ ] This form is complete
Signed: _____________

### 2.3 Memory Test

Form B: Memory Test - Subject #__
Aman Bhargava's AP Capstone Research Project: To what extent can self-reported interest be predicted using EEG and machine learning?

Outliers Excerpt:

1. What was the question that psychologists have been studying?
   ____________________________________________
   ____________________________________________
   __________________________
2. Where was the university at which the study was conducted? (Circle one)
   a. Warsaw
   b. Berlin
   c. London
   d. Paris
3. What was the main conclusion of the study?
   ____________________________________________
   ____________________________________________
   ____________________________________
4. How many hours must a person practice to become 'great' at something, according to the article?
   ____________________________________________
   _____________
5. What was the name of the psychologist who conducted the study at the music academy?

________________________________________
____________

Wikipedia Article on Mathematics:

1. Which language did the word 'mathematics' come from? (Circle one)
   a. Latin
   b. Arabic
   c. Greek
   d. Persian
2. How do mathematicians determine the validity of a conjecture?

. Mathematical argument
a. Mathematical theorizing
b. Mathematical proof
c. Mathematical debate

3. Where did rigorous mathematical arguments first appear?

. Greece
a. Rome
b. Persia
c. India

4. Who said, "The universe cannot be read until we have learned the language and become familiar with the characters in which it is written"?

. Galileo
a. Aristotle
b. Euphrates
c. Archimedes

5.

Wikipedia Article on Benjamin Franklin:

1. Please name two inventions by Benjamin Franklin that were discussed or mentioned in the article.

________________________________________
________________________________________
_________________________________

2. In which state did Franklin become a successful newspaper editor and printer?

a. Boston
b. London
c. Waterloo
d. Philadelphia

3. From which country did Franklin secure support from for the American war of independence?

. Scotland
a. France
b. Germany
c. Italy

4. What was the first name of Benjamin Franklin's father?

. Josiah
a. Elisiah
b. Benjamin I
c. John

5. Which present-day university did Benjamin Franklin pioneer?

. Princeton
a. Harvard
b. Yale
c. University of Pennsylvania

### 2.4 Interest Levels

Form C: Memory Test - Subject #__
Aman Bhargava's AP Capstone Research Project: To what extent can self-reported interest be predicted using EEG and machine learning?

Reading #1: ____________________________

*Deeply Uninteresting* 1 2 3 4 5 *Extremely Interesting*

*Binary Value:* 1 / 0

Reading #2: ____________________________

*Deeply Uninteresting* 1 2 3 4 5 *Extremely Interesting*

*Binary Value:* 1 / 0

Reading #3: ____________________________

*Deeply Uninteresting* 1 2 3 4 5 *Extremely Interesting*

*Binary Value:* 1 / 0

---

This form is fully completed [ ]
Signed: _______________________

## APPENDIX 3: MACHINE LEARNING ALGORITHM PARAMETERS

### 3.1 Decision Forest

Binary Gemini Decision Forest Classifier

Settings

| Setting | Value |
|---|---|
| Ensemble Element Count | 8 |
| Max Depth | 32 |
| Random Split Count | 128 |
| Min Leaf Sample Count | 1 |
| Class Count | 2 |
| Resampling Method | Bagging |
| Random Number Seed | 5 |
| Allow Unknown Levels | True |

### 3.2 Logistic Regression

## Logistic Regression Classifier

### Settings

| Setting | Value |
|---|---|
| Optimization Tolerance | 1E-07 |
| L1 Weight | 1 |
| L2 Weight | 1 |
| Memory Size | 20 |
| Quiet | True |
| Use Threads | True |
| Allow Unknown Levels | True |
| Random Number Seed | |

### 3.3 Support Vector Machine

## Support Vector Machine Classifier

### Settings

| Setting | Value |
|---|---|
| Lambda | 0.001 |
| Num Iterations | 1 |
| Normalize Features | True |
| Perform Projection | False |
| Allow Unknown Levels | True |
| Random Number Seed | |

### 3.4 Two-Class Neural Net

## APPENDIX 4: FEATURE EXTRACTION CODE

### 4.1 Main Java Separator Code

```
import java.io.File;
import java.io.FileNotFoundException;
import java.io.PrintWriter;
import java.io.UnsupportedEncodingException;
import java.util.ArrayList;
import java.util.Scanner;
public class CSV_Splitter {

  /**
   * This class will split CSV data into 3 files.
   * They will be split based on the numbers in MARKER_COLUMN = 19;
```

### Settings

| Setting | Value |
|---|---|
| Loss Function | CrossEntropy |
| Learning Rate | 0.1 |
| Number Of Iterations | 1000 |
| Is Initialized From String | False |
| Is Classification | False |
| Initial Weights Diameter | 0.1 |
| Momentum | 0 |
| Neural Network Definition | |
| Data Normalizer Type | MinMax |
| Number Of Input Features | |
| Number Of Hidden Nodes | System.Collections.Generic.List`1[System.Int32] |
| Number Of Output Classes | |
| Shuffle | True |
| Allow Unknown Levels | True |
| Random Number Seed | |

```
   * These are the things that need to happen for this method:
   * 1. Find indexes of the marker numbers
   * 2. Store those in a 2x6 array
   * 3. Call the outputListArr method on the array multiple times to export
   *
   * I need to iterate through:
   * 1. Each test subject (1, 2, 3.csv)
   * 2. Each reading for each test subject (1, 2, 3) (done manually, not a formal loop)
   * @param args
   * @throws FileNotFoundException
   * @throws UnsupportedEncodingException
   */
  public static void main(String[] args) throws FileNotFoundException, UnsupportedEncodingException {
          // TODO Auto-generated method stub

          for(int i = 1; i <= 10; i++) {
                  String inName = "../DATA/Originals/"+i+".csv";
```

```
                        System.out.println(inName);

                        ArrayList<String[]> curCSV =
        CSV_Test.getList(inName);

                        ArrayList<int[]> markerColumn =
        CSV_Test.getMarkerVals(curCSV);

                        for(int j = 1; j <=3; j++) {
                                  String outName = i + "-
        " + j;
                                  int startInd = j*2-1;
                                  int endInd = start-
        Ind+1;

                                  CSV_Write.output-
        ListArr(curCSV, outName, markerColumn.get(startInd-
        1)[1], markerColumn.get(endInd-1)[1]);
                        }
              }
    }
}
```

## 4.2 Main Octave FFT Code

```
% We are trying to extract FFT from each column in
the matrix we are given.

for u = 1:10,
          for j = 1:3,
                    fileName = strcat(mat2str(u),"-
",mat2str(j),".csv");

                    x = csvread(fileName);
                    x = x(:, 3:16);

                    superMatrix = [];

                    for i = 1:size(x)(2), % going
through the 14 channels on the EEG
                              tmp = x(:, i); %this
is a single column.
                              a = fft(tmp);
                              b = abs(a);
                              b = b(1:size(b)/2);
                              superMatrix = [super-
Matrix, b];
                    endfor;
                    display(fileName);
                    display(size(superMatrix));
                    outName =
strcat("FFT/",mat2str(u),"-",mat2str(j),"-
RFFT.csv");

                    dlmwrite(outName, superMatrix);
% output b to csv for later processing in Java
          endfor;
endfor;
```